Enclosure

Request for Data from Epidemiological Studies by Independent Research Institutions

Your letter of March 4, 2013 requested access to data analyzed in several studies of long-term exposure to $PM_{2.5}$ based on the American Cancer Society (ACS) cohort and the Harvard Six Cities cohort, as well as a study of long-term exposure to ozone based on the ACS cohort. As you note in your letter, the EPA has received and responded to previous requests to provide research data from these epidemiological studies that were partially funded by EPA grant funds. We provide more information regarding these datasets, the legal framework, and the EPA's responses below.

The EPA is committed to compliance with the requirements of the Shelby Amendment (Public Law 105-277) and to transparency with regard to the scientific bases of agency decision making, and to increasing access to federally-sponsored scientific research as outlined in the recent memorandum[1] from the Office of Science and Technology Policy (OSTP) in the Executive Office of the President. Both the Shelby Amendment and the OSTP memorandum recognize that increasing access to federally funded, scientific data must be balanced with requirements to protect the research subject's privacy.

The underlying data you are requesting for each epidemiological study consist of three distinct datasets, which the researchers link together in order to estimate the relative risks of exposure to air pollution: (1) air quality data; (2) health event data, which in these studies are data from the National Death Index (NDI); and (3) individual health data that are gathered through detailed questionnaires[2] completed for each study participant in the cohort. The complete, linked set of data underlying these studies is held by the scientific researchers that conducted the relevant research, not the EPA. Further, the availability of some of these datasets is subject to certain non-disclosure protections, which we describe below.

These studies are large epidemiological research projects that received funding for different components of data collection and analysis from a number of different sources, including the EPA, other federal agencies, and non-federal sources. The Shelby Amendment directs the Office of Management and Budget (OMB) "to require Federal awarding agencies to ensure that all data produced under an award [of federal funds] will be made available to the public through the procedures established under the Freedom of Information Act."[3] The OMB implemented the Shelby Amendment by amending the government-wide regulations applicable to grant awards

[1] Holdren, John P. *Memorandum for the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research.* Office of Science and Technology Policy. Executive Office of the President. February 22, 2013.
http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

[2] The questionnaires for these studies requested very detailed personal information, including questions on residential location, age, race, educational attainment, body mass index, alcohol consumption, smoking history, occupational exposure to pollution, and medical history.

[3] "Provided further, that the Director of OMB amends Section __.36 of OMB Circular A-110 to require Federal awarding agencies to ensure that all data produced under an award will be made available to the public through the procedures established under the Freedom of Information Act." (Public Law 105-277).

for research grants.[4] The EPA has adopted this provision into the Agency's grant regulations.[5] These regulations require the EPA to request, and award recipients to provide, "research data" produced with the support of federal funding. Under the regulations, "research data" does not include "medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study".[6] The exclusion of this type of information from disclosure is consistent with the personal privacy exemption contained in the Freedom of Information Act (FOIA).[7]

In addition to non-disclosure protections under the Shelby Amendment and the FOIA, the health event data from NDI[8] are also governed by the Public Health Service Act (PHSA), which prohibits the release of information that may identify the person or institution supplying the information. Specifically, the PHSA provides that data collected by the National Center for Health Statistics – which includes the individual-level health event data from the NDI used in the relevant epidemiological studies – may not be used for any purpose other than the purpose for which it was supplied and not be released if the information could be personally identifying.[9] Furthermore, the PHSA does not allow release of such information unless such establishment or individual has consented in its publication or release in other form. To obtain data in the NDI, researchers sign a confidentiality agreement promising not to publish or release the data in any form to any party if a particular individual (or establishment) is identifiable.[10]

---

[4] OMB Circular A-110, which is codified at 2 CFR 215.36: These regulations provide that "in response to a Freedom of Information Act (FOIA) request for research data relating to published research findings produced under an award that were used by the Federal Government in developing an agency action that has the force and effect of law, the Federal awarding agency shall request, and the recipient shall provide, within a reasonable time, the research data so that they can be made available to the public through the procedures established under the FOIA."

[5] 40 CFR 30.36

[6] 2 CFR 215.36(d)(2)(i) and 40 CFR 30.36(d)(2)(i)(B).

[7] The FOIA requires the EPA to release such information upon request unless the information is protected from disclosure under its FOIA exemptions. 5 U.S.C. § 552(b). Exemption 6 of FOIA protects "personnel and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy." 5 U.S.C. § 552(b)(6).

[8] The NDI is part of the National Center for Health Statistics at the Centers for Disease Control and Prevention. The NDI is a central computerized index of death record information that was established as a resource for epidemiologists and other health and medical investigators.

[9] Public Health Service Act (42 U.S.C. § 242m(d)): "No information, if an establishment or person supplying the information or described in it is identifiable, obtained in the course of activities undertaken or supported under section 242b, 242k, or 242l of this title may be used for any purpose other than the purpose for which it was supplied unless such establishment or person has consented (as determined under regulations of the Secretary) to its use for such other purpose; and in the case of information obtained in the course of health statistical or epidemiological activities under section 242b or 242k of this title, such information may not be published or released in other form if the particular establishment or person supplying the information or described in it is identifiable unless such establishment or person has consented (as determined under regulations of the Secretary) to its publication or release in other form."

[10] To use the NDI data, researchers agree to a Data Use Agreement stating that they will not link these data with individually identifiable records from any other National Center for Health Statistics (NCHS) or non-NCHS data set. http://www.cdc.gov/nchs/data_access/restrictions.htm

In response to a letter from Representative Harris on September 22, 2011, the EPA requested the "research data" (as defined by applicable EPA regulations described above) funded by EPA grant funds from Harvard University and New York University related to two epidemiological studies.[11,12] Both institutions provided the air quality data in response to this request. Harvard University also provided health event data from the NDI. Neither institution provided the individual health questionnaire data to EPA. Prior to disseminating the NDI data provided by Harvard University, the EPA coordinated with the Centers for Disease Control and Prevention to ensure that the data did not identify the particular establishment or individual supplying the information. In June 2012, the EPA provided to Representative Harris all of the data received from Harvard University and New York University.

The EPA recognizes that the data provided in response to the request from Representative Harris are not sufficient in themselves to replicate the analyses in the epidemiological studies for two reasons. First, these cohort analyses on premature mortality relied on linking medical and demographic information with air quality data. The combination of these data could identify specific individuals and thus could not be released in its original format without consent of the study participants. Second, as noted above, it appears that some of the underlying data collection may have been funded by other government entities. For these reasons, composing a data set sufficient to even generally replicate the published analyses to which you refer is a complicated undertaking requiring the input of several funding agencies, awardees, and the resources of federal (or other) non-disclosure boards to ensure that the data are not identifiable.

In the past, the Health Effects Institute (HEI) entered confidentiality agreements with the researchers to have access to the data in order to conduct a full reanalysis of two studies of these cohorts,[13] which HEI completed in 2000.[14] In order to access the private medical information from the original investigators, HEI guaranteed that confidentiality that had been provided to

---

[11] Pope, C.A., III, R.T. Burnett, M.J. Thun, E.E. Calle, D. Krewski, K. Ito, and G.D. Thurston. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution." *Journal of the American Medical Association* 287:1132-1141.

[12] Laden, F., J. Schwartz, F.E. Speizer, and D.W. Dockery. 2006. Reduction in Fine Particulate Air Pollution and Mortality. *American Journal of Respiratory and Critical Care Medicine.* 173: 667-672.

[13] HEI Statement, 2000, p. i: "Both of these studies came under intense scrutiny in 1997 when the EPA used the results to support new National Ambient Air Quality Standards for fine particles and to maintain the standards for particles less than 10 μm in median aerodynamic diameter (PM_{10}) already in effect. Members of Congress and industry, the scientific community and others interested in regulation of air quality scrutinized the studies' methods and their results. Some insisted that any data generated using federal funding should be made public. Others argued that these data had been gathered with assurances of confidentiality for the individuals who had agreed to participate and that the concept of public access to federally funded data did not take into account the intellectual property rights of the investigators and their supporting institutions. To address the public controversy, Harvard University and the ACS requested that the Health Effects Institute organize an independent reanalysis of the data from these studies. Both institutions agreed to provide access to their data to a team of analysts to be selected by HEI through a competitive process." http://pubs.healtheffects.org/getfile.php?u=271

[14] Krewski, D., R.T. Burnett, M.S. Goldbert, K. Hoover, J. Siemiatycki, M. Jerrett, M. Abrahamowicz, and W.H. White. 2000. "Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality." Special Report to the Health Effects Institute. Cambridge MA. July. http://pubs.healtheffects.org/getfile.php?u=274

study participants by the original investigators would be fully respected by the reanalysis team.[15] It may be possible for other researchers to contact the original researchers and gain access to the data by entering into similar confidentiality agreements.

We note that in setting the National Ambient Air Quality Standards (NAAQS) and in assessing health benefits anticipated from air pollution regulations, the EPA relies on the scientific studies that were published in the peer-reviewed literature rather than the underlying data that consist of private medical information.[16] The EPA provides the information used in regulatory decisions, including the epidemiological studies, in the publicly available docket accompanying each rulemaking.

Your letter of March 4, 2013, requested access to the data analyzed in the $PM_{2.5}$ studies based on the ACS cohort and the Six Cities cohort cited in the final PM NAAQS rule including newer studies documenting extended analyses of these two cohorts that included additional years of follow-up.[17,18] and the underlying data from the long-term ozone study based on the ACS cohort.[19] In response to the request from Representative Harris noted above, the EPA has already provided all of the information funded by EPA grant funds for the earlier studies that the researchers provided to the EPA. We are enclosing these data in this response. In addition, the EPA is willing to submit a new request to the research institutions for the research data corresponding to the additional years of follow-up in the newer studies cited in your letter of March 4, 2013. We note that, because of the limitations on disclosure of personal private information, the information that could be disclosed for both the older and the newer epidemiological studies would be insufficient to replicate the analyses in those studies. Again, we would be pleased to meet with your staff to discuss these matters in greater detail. Additionally, as noted above, there remains the possibility that other independent researchers

---

[15] HEI preface, 2000, p.2: "Both conducting the work and reporting the results would be as open and public as possible. The guarantees of confidentiality that had been provided to study participants by the Six Cities Study and the ACS Study Original Investigators would be fully respected by the Reanalysis Team. Beyond this, any methods used, analyses undertaken, and results produced would be completely and publicly described." http://pubs.healtheffects.org/getfile.php?u=273

[16] See, for example the 1997 PM NAAQS, 62 Fed. Reg. at 38691/1: "EPA did not rely upon the raw health data supporting the Dockery and Pope studies; it relied instead upon the studies themselves. These studies may properly be considered 'data' The EPA has never had the raw data in its possession; thus EPA has neither reviewed it nor had an opportunity to place it in the docket. The EPA did rely on the studies and these studies are included in the docket and are available for public review. Because EPA neither reviewed nor relied upon the raw data, there is no obligation to obtain it or to make it available."

[17] Lepeule J, Laden F, Dockery D, Schwartz J 2012. "Chronic Exposure to Fine Particles and Mortality: An Extended Follow-Up of the Harvard Six Cities Study from 1974 to 2009." *Environ Health Perspect.* Jul;120(7):965-70.

[18] Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi, Y, et al. 2009. "Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality." *HEI Research Report, 140,* Health Effects Institute, Boston, MA.

Lepeule J, Laden F, Dockery D, Schwartz J 2012. "Chronic Exposure to Fine Particles and Mortality: An Extended Follow-Up of the Harvard Six Cities Study from 1974 to 2009." *Environ Health Perspect.* Jul;120(7):965-70.

[19] Jerrett, M; Burnett, RT; Pope, CA, III; Ito, K; Thurston, G; Krewski, D; Shi, Y; Calle, E; Thun, M. (2009). "Long-term ozone exposure and mortality." *N Engl J Med* 360: 1085-1095.

could contact the original researchers and gain access to the data by entering into confidentiality agreements similar to those that permitted the HEI reanalysis.

Questions on Ozone Integrated Science Assessment

Your letter also questioned the EPA's interpretation of specific epidemiological studies and causality determinations discussed in the Integrated Science Assessment (ISA) for Ozone, which was finalized in February 2013.[20] In developing an ISA, the EPA uses a formal causal framework that provides a consistent and transparent basis for integration of scientific evidence and evaluation of the causal nature of air pollution-related health effects. This approach has been reviewed and endorsed by the Clean Air Scientific Advisory Committee (CASAC).[21] This framework employs a five-level hierarchy that classifies the overall weight of evidence and causality using the following categorizations: causal relationship; likely to be a causal relationship; suggestive of a causal relationship; inadequate to infer a causal relationship; and not likely to be a causal relationship. Pursuant to this framework, in order to reach a determination that the weight of scientific evidence is suggestive of a causal relationship, the evidence should include "at least one high-quality epidemiologic study show[ing] an association with a given health outcome."[22]

The previous scientific assessment for ozone[23] in 2006 concluded that an insufficient amount of evidence existed to suggest a causal relationship between chronic ozone exposure and increased risk of mortality in humans. However, two recent studies[24,25] provided new evidence for the 2013 assessment. This new evidence is consistent and coherent with the evidence from epidemiological, controlled human exposure, and animal toxicological studies for the effects of short- and long-term exposure to ozone on respiratory effects. The current body of evidence, including these two high-quality, peer-reviewed studies that observed associations between long-term exposure to ozone and mortality, is suggestive of a causal relationship between long-term exposure to ozone and total mortality.

---

[20] U.S. EPA (U.S. Environmental Protection Agency). 2013. *Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report).* U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-10/076F. http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492

[21] CASAC review (Samet, 2009). Letter to Administrator Jackson in response to review of Integrated Science Assessment for Particulate Matter (Second External review Draft, July 2009): "As mentioned in its comments on the charge questions, CASAC also commends EPA for the continued evolution of the process for evidence evaluation. The five-level classification of strength of evidence for causal inference has been systematically applied; this approach has provided transparency and a clear statement of the level of confidence with regard to causation, and we recommend its continued use in future ISAs." http://yosemite.epa.gov/sab/sabproduct.nsf/264cb1227d55e02c85257402007446a4/151B1F83B023145585257678006836B9/$File/EPA-CASAC-10-001-unsigned.pdf

[22] Ozone ISA (p. lxvii).

[23] U.S. EPA (U.S. Environmental Protection Agency). 2006. *Air quality criteria for ozone and related photochemical oxidants [EPA Report].* (EPA/600/R-05/004AF). Research Triangle Park, NC. http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=149923

[24] Jerrett et al. (2009).

[25] Zanobetti, A; Schwartz, J. 2011. "Ozone and survival in four cohorts with potentially predisposing diseases." *Am J Respir Crit Care Med* 184: 836-841.

Your letter stated that 11 earlier studies did not find statistically significant associations between long-term exposure to ozone and mortality and that the EPA selectively relied on the one positive study to support the causality determination of "suggestive." A key explanation for the lack of associations found in most of these earlier studies is that they did not specifically assess respiratory mortality.[26] However, unlike the earlier studies, Jerrett et al. (2009) did specifically evaluate respiratory mortality and found a statistically significant association. This finding is consistent with other studies finding associations with respiratory effects (e.g., morbidity and mortality). Because of the strength of the evidence between ozone exposure and respiratory effects, it is reasonable to find associations between long-term exposure to ozone and respiratory mortality but not other sources of mortality (e.g., all-cause, cardiovascular, and cardiopulmonary). Consequently, EPA concluded that there is sufficient evidence "suggestive of a causal relationship" between long-term exposure to ozone and respiratory mortality consistent with the formal causal framework, which has been reviewed and endorsed by the CASAC.

[26] Abbey et al. (1999) is the only one of the cited studies other than Jerrett et al. (2009) that assesses respiratory mortality, finding no association. In addition, your letter does not identify several additional studies that found associations between long-term exposure to ozone and all-cause mortality that were identified in the ozone ISA (Lipfert et al., 2000, 2006; Smith et al., 2009).

Abbey et al. 1999. "Long-term inhalable particles and other air pollutants related to mortality in nonsmokers." *Am. J. Rep. Cril. Care Med* 159: 373-382.

Lipfert, FW; Perry, HM, Jr; Miller, JP; Baty, JD; Wyzga, RE; Carmody, SE. 2000. The Washington University-EPRI veterans' cohort mortality study: Preliminary results. *Inhal Toxicol* 4: 41-73.

Lipfert, FW; Perry, HM, Jr; Miller, JP; Baty, JD; Wyzga, RE; Carmody, SE. 2003. Air pollution, blood pressure, and their long-term associations with mortality. *Inhal Toxicol* 15: 493-512.

Smith, KR; Jerrett, M; Anderson, HR; Burnett, RT; Stone, V; Derwent, R; Atkinson, RW; Cohen, A; Shonkoff, SB; Krewski, D; Pope, CA, III; Thun, MJ; Thurston, G. 2009. "Public health benefits of strategies to reduce greenhouse-gas emissions: Health implications of short-lived greenhouse pollutants." *Lancet* 374: 2091-2103.